

**Pamela Carson**  
**Concordia University Library, Montreal, Quebec, Canada**

**Francisco Berrizbeitia**  
**Concordia University Library, Montreal, Quebec, Canada**

# **SUPPORTING GLOBAL EQUITY IN DOCUMENTARY HERITAGE WITH A FREE, MINIMAL COMPUTING SOLUTION FOR CREATING SUSTAINABLE AND ACCESSIBLE DIGITAL EXHIBITIONS: THE CASE OF ZULIMA**

## **Abstract**

The Global South's documentary heritage is less online than the Global North's and tools and methods are needed for enhancing the discoverability of collections where resources are constrained. The Zulima collection, late 19<sup>th</sup>- and early 20<sup>th</sup>-century documents pertaining to an important female Venezuelan author and playwright is a collection at risk of being lost as there are currently no memory institutions in Venezuela that can safely keep it. This project proposes a way to expand on existing metadata management software, SWALLOW, minimal web templates, a data export to Wikidata, and free asset hosting, to make this collection available online and accessible for future researchers.

## **Introduction**

### ***Background***

Over the last two decades, Dr. Mariana Libertad Suárez, a Venezuelan university professor, has dedicated herself to researching Latin American women writers from the nineteenth and early twentieth century who are absent from the literary canon. She has amassed a significant amount of information about early women writers from Venezuela, including primary-source out-of-copyright texts and manuscripts, correspondence, and images, as well as research on these writers. This collection includes priceless items about Zulima (pseudonym for Lina López de Aramburu), the first female Venezuelan playwright and one of the first female Venezuelan authors.

Libertad Suárez left Venezuela years ago due to the ongoing socio-political and economic crisis there and has brought this archive with her to Spain where she currently teaches at the Pontificia Universidad Complutense de Madrid.

Neely-Cohen (2025) wrote that if someone wants to store something for 100 years, the best way would be to keep it within a social or governmental institution that is likely to provide care. The

current situation in Venezuela is dire, with hyperinflation, poverty and widespread hunger, crime and corruption. Over 7.7 million Venezuelans have left their country in recent years since the beginning of the Bolivarian Revolution, contributing to the largest recorded refugee crisis in the Americas (UNHCR, 2025). Early in 2025, Nicolás Maduro was inaugurated for the third time following a disputed election. U.S., Canada and Europe have increased sanctions on Venezuelan officials, and the U.S. is offering \$25 million for Maduro's arrest (Spetalnick & Parraga, 2025).

In such an environment, digital preservation of cultural heritage is not a priority. In fact, regions with centralized power structures such as Venezuela with its single-party rule and authoritarian regime, authority is maintained by controlling information and history has seen many examples of the intentional destruction of artifacts and records, censorship of the media, and the silencing of critics (Budd, 2025).

In Venezuela, there is no public infrastructure available for hosting digital archives. While similar projects such as Colombia's *Soledad Acosta de Samper Digital Library*<sup>1</sup> have been created in Latin American countries with better public infrastructure, Venezuelan cultural heritage is at risk of being forgotten because of barriers to digitization. As the popularity of AI-powered search increases and sources are obscured, it is increasingly important to get Venezuelan cultural data online so that it can be included and not forgotten.

### ***Problem statement***

Worldwide, memory institutions continue to work on making their collections available online. While the Global North has both national and international repositories for digital cultural heritage (e.g., *Europeana*, *Canadiana*, the *Digital Public Library of America*, among others) and countless memory institutions – galleries, museums, archives and libraries – the Global South is comparatively underserved. Some of the Global South's heritage may be available through the Global North's platforms because archives were expatriated from source locations, and this is controversial (Lovering, 2009) and repatriating archives is something institutions have started to investigate (Christen & Pugh, 2011). The Zulima collection is one example of digital cultural heritage at risk and surely there are many more examples of such collections. First, there is no memory institution in Venezuela to safely hold the collection. The software generally used to create digital collections either is expensive or requires technical expertise to set up and maintain. In order to make this collection accessible to future researchers it needs to be online, preferably in a structured data format that is machine readable, and which offers the highest chance of being discoverable.

### ***Objectives***

---

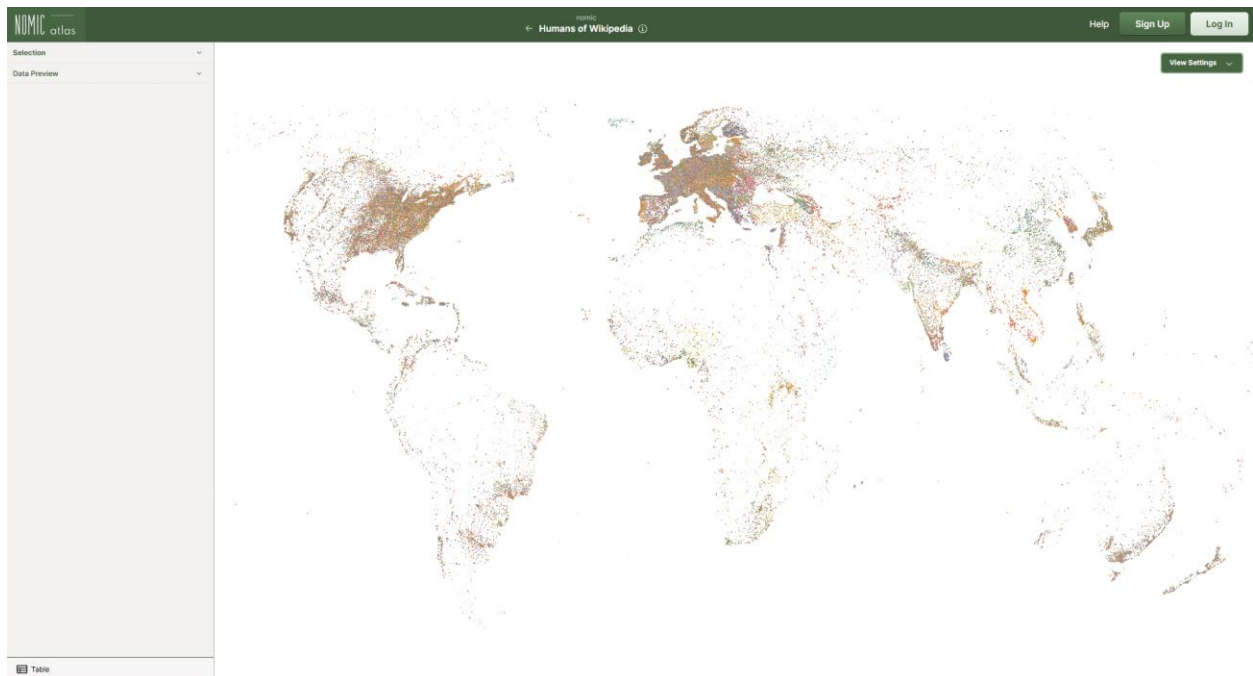
<sup>1</sup> <https://soledadacosta.uniandes.edu.co/>

The objective of this project is to make the Zulima collection, and other collections like it, available online, discoverable and fully accessible in a way that requires no cost, minimal labour, and limited ongoing maintenance.

1. To expand on the existing SWALLOW software (Camlot, Neugebauer, & Berrizbeitia, 2020) to make it suitable for cataloguing collections with similar constraints (i.e., create a portable application that supports multiple interface languages).
2. To support the creation of minimal digital collection websites that use only JavaScript, local data, standards-based HTML and CSS to create a self-contained, sustainable collection that can be hosted anywhere.
3. To create machine-readable, linked data from the cataloguing data and make it available on Wikidata for both creating Wikipedia pages and for data harvesting.
4. To find suitable and sustainable options for hosting digital assets from the collection (Internet Archive, Flickr Commons, Wikimedia Commons, among others).

### ***Significance***

Ferdinand (2000) noted that, in the early days of the internet, there were predictions that it would “completely revolutionize government and democracy, even that the outcome [would] be a new wave of democratization world-wide, as authoritarian regimes [would] find it difficult to survive” (p. 1), but that sadly, the greatest impact at that time was instead on business and commerce rather than democracy. In Figure 1, Iaouenan et al. (2021) visualized all the notable people in Wikipedia and mapped them according to birthplace. The data is dense in Europe and the United States, and remarkably sparse in South America and Africa. Cultural representation is unequal and uneven globally. In the context of digital humanities, Risam (2019) encourages the creation of “new tools and methods with epistemological and ontological roots beyond the Global North” (p. 25).



*Figure 1 Birthplace locations of notable people in Wikipedia (3500 BC to 2018 AD). Note the imbalance between the Global North and the Global South (Iaouenan et al., 2021).*

## Literature review

*The digital cultural record, in its gaps, omissions, and rehearsal of colonial dynamics from the analog cultural record, also has a chronic problem of epistemic violence, which contravenes cultural survival for communities whose languages are underrepresented, histories are suppressed, and stories are untold (Risam, 2019, p. 4)*

UNESCO, the United Nations Educational, Scientific and Cultural Organization, has concerned itself with documentary heritage for decades. In 2015, it highlighted the importance of documentary heritage on many fronts and encouraged member states to safeguard it. UNESCO (2021b) established the PERSIST project (Platform to Enhance the sustainability of the Information Society Transglobally) with cooperation from international memory institutions (IFLA and ICOM), government agencies and the ICT industry. However, there is no actual software platform – the project was a series of consultations which produced policies and recommendations. UNESCO’s Information For All Programme (IFAP) Strategic Plan 2023-2029 mentions as objectives “promote and widen access through the organization, digitization and preservation of information,” but frames the preservation of information as part of the “digital transformation of memory institutions” where it will provide “metrics, policy guidance, capacity development and exchanges of experiences” (p. 8) (UNESCO, 2023). Again, within the confines of memory institutions, a project output is to help safeguard documents held by museums, libraries and archives. While UNESCO has drawn attention to this area, no tangible assistance is apparently available for people wanting to catalogue and/or digitize collections aside from a

handful of funded projects (UNESCO, 2024) and guidance on low-tech digitization (UNESCO, 2021b).

While infrastructure such as LOCKSS and CLOCKSS<sup>2</sup> exists for digital preservation at well-funded memory institutions that can absorb ongoing costs, this infrastructure is out of reach for Venezuelan institutions. In a zero-budget scenario, an approach that assumes no ongoing costs is required. In addition, even if state-funded memory institutions were providing access to this infrastructure, it is likely in a country like Venezuela that certain items could be censored. Open-source digital preservation solutions are ideal to avoid having data locked in by vendors or in proprietary formats and to minimize the financial cost. Omeka<sup>3</sup>, an open-source software for digital collections, is commonly used for community archives and smaller projects, but it either requires technical expertise to maintain this PHP-based software with multiple external dependencies as well as space to host it or an ongoing subscription for a hosted version. CONTENTdm<sup>4</sup>, which costs thousands of dollars per year, is commonly used by institutions, as well as free, open-source software such as Collective Access<sup>5</sup>, Collection Space<sup>6</sup>, which require technical expertise and resources to host.

## **Methodology**

There are multiple examples of software helping empower marginalized communities. This type of software has been called ICT4D, information and communication tools for development, and has been proposed by Frings-Hessami and Oliver (2023) as something that can be combined with archival science.

## ***Project design***

For the [SpokenWeb](#), a related SSHRC-funded digital humanities project led by Concordia English professor Jason Camlot, Francisco Berrizbeitia and Tomasz Neugebauer developed SWALLOW<sup>7</sup>, a lean, open-source document-oriented database for ingesting metadata. SWALLOW will be re-used to manage metadata for the Zulima collection.

Once metadata is entered into SWALLOW by either Venezuelan graduate students working under Professor Suárez or graduate students in Canada, digital assets will be uploaded to the [Internet Archive](#) or another suitable repository for public domain archival material such as [Flickr Commons](#). The digital assets (texts and images) from the collection are all in the public domain now.

---

<sup>2</sup> <https://www.lockss.org/> and <https://clockss.org/>

<sup>3</sup> <https://omeka.org/>

<sup>4</sup> <https://www.oclc.org/en/contentdm.html>

<sup>5</sup> <https://collectiveaccess.org/>

<sup>6</sup> <https://collectionspace.org/>

<sup>7</sup> SWALLOW source code: <https://github.com/lib-development-concordiaU/swallow2> and SWALLOW documentation: <https://github.com/lib-development-concordiaU/swallow2/wiki>

Also, a free website will be created on [GitHub Pages](#) for the collection.

## ***Project plan***

### *Pre-project*

Before starting the project, we will measure the visibility of Zulima and her works online in search engines, chatbots, in multiple languages as a pre-test. A set of standard questions to be asked again at the end of the project to see the impact of the project.

We are also applying for a SSHRC Connection Grant for funding to hire students to assist in software development, writing documentation and cataloguing.

### *Part 1A: Software development*

#### *Portable application*

Building on the current version of SWALLOW, we will develop a portable version of the software so that it can be run as a standalone software on any compatible computer without needing to be installed. We will containerize Swallow with Docker<sup>8</sup>, i.e., modify the software so that all its components are bundled into a single image that can be run on any operating system.

#### *Software localization*

Swallow's interface is currently only available in English. We will make it localizable, i.e., build the capacity to have the software in multiple languages, and translate relevant parts of the software and its schema labels into Spanish, French, and possibly other languages. As we develop SWALLOW for use for creating digital archives in resource-constrained contexts, we also want to make it as accessible as possible by providing the software in multiple languages (Khan & Ullah, 2024). Many languages are spoken by librarians and staff at the Concordia Library including French, Spanish, Polish, Cantonese, Amharic, and we would like to invite our colleagues to contribute translations for the software.

#### *Schemas*

To output structured and consistent data from SWALLOW to both WikiData and the digital exhibition website, the appropriate schemas for each item type will be selected from Schema.org, translated, and pre-loaded into the software.

#### *Documentation*

We will hire a Library Studies student to build on existing documentation and have this translated also.

### *Part 1B: Cataloguing the Zulima collection*

---

<sup>8</sup> <https://www.docker.com/>

Concurrent with software development, cataloguing work can start. Libertad Suárez will work with a PhD candidate student to catalogue the collection in SWALLOW and digitize selected assets.

### ***Part 2A: Populating WikiData***

Additional work will need to be done to find the WikiData equivalents to crosswalk data.

### ***Part 2B: Creating the website on GitHub Pages***

Another output of the cataloguing is to create a minimal website that can be hosted for free on GitHub Pages, but in theory could be hosted anywhere.

### ***Part 3: Asset management***

Optimized images, aiming for smallest file size, smallest amount of work, with the most legibility for images of text.

Since a requirement of the project is to host the digital exhibition website for free, we need to find a long-term storage space for digitized assets. While research data management workflows and storage solutions for the sciences are common, the management of humanities research data is less established. Possibilities exist with Wikimedia Commons, Flickr Commons, Project Gutenberg, and the Internet Archive.

### **Expected outcomes**

Our hypothesis is that cataloguing the Zulima collection, creating a digital collection website and exporting the data to Wikidata will increase the visibility and discoverability of this author's works online. We will test this by re-asking the pre-test questions in Google and chatbots to see if results include Zulima. Over time, we would like to see if citations of works in the collection increase and if more information is added by others to Wikipedia once the data seeds are planted.

## References

- Berrizbeitia, F., & Neugebauer, T. (2023) A linked data approach to SpokenWeb collections. In: SpokenWeb 2023 Institute, Thu, May 4, 2023 - Fri, May 5, 2023, Edmonton, Alberta. <https://spectrum.library.concordia.ca/id/eprint/992255/>
- Berrizbeitia, F. (2022) An interactive AI-based approach to semantic annotations for the SpokenWeb archive. In: 20th Annual Research Forum, 26 April 2022, Montreal. <https://spectrum.library.concordia.ca/id/eprint/992198/>
- Budd, J. M. (2025). Authoritarian/totalitarian censorship, and an ethical response. *The Library Quarterly*, 95(1), 42-57. <https://doi.org/10.1086/733175>
- Camlot, J., Neugebauer, T., & Berrizbeitia, F. (2020). [Dynamic systems for humanities audio collections: The theory and rationale of Swallow](#). In: DH 2020 (Digital Humanities 2020 Virtual Conference), July 23, 2020, Ottawa, Canada.
- Christen, K., & Pugh, M. (2011). Opening archives: Respectful repatriation. *The American Archivist*, 74(1), 185-210. <https://doi.org/10.17723/aarc.74.1.4233nv6nv6428521>
- Eze Asogwa, B. (2011). Digitization of archival collections in Africa for scholarly communication: Issues, strategies, and challenges. *Library Philosophy and Practice*, 651. <https://digitalcommons.unl.edu/libphilprac/651>
- Ferdinand, P. (2000). The internet, democracy and democratization. *Democratization*, 7(1), 1-17. <https://doi.org/10.1080/13510340008403642>
- Frings-Hessami, V., & Oliver, G. (2023). Accessing and preserving information: Combining ICT4D and archival science to empower marginalized communities. *Journal of the Association for Information Science and Technology*, 74, 1350–1364. <https://doi.org/10.1002/asi.24702>
- Iaouenan, M., Eyméoud, J. B., Gergaud, O., Bhargava, P., Plique, G., & Wasmer, E. (2021). A cross-verified database of notable people, 3500bc-2018ad. *CEPR Discussion Paper*, DP15852. <https://ssrn.com/abstract=3795248>
- Khan, M. T., & Ullah, M. (2024). Enhancing accessibility of archives of developing countries: Strategies for inclusive information access in Pakistani context. *Proceedings of the Association for Information Science and Technology*, 61(1), 968–970. <https://doi.org/10.1002/pra2.1156>
- Lovering, T. (2009). Expatriate archives. *Archives: The Journal of the British Records Association*, 34(121). <https://doi.org/10.3828/archives.2009.10>
- Melanson, J. (2020). Towards epistemic justice in the archives: The implications of Miranda Fricker's philosophy for archival reference. *Emerging Library & Information Perspectives*, 3, 89–112. <https://doi.org/10.5206/elip.v3i1.8617>
- Neely-Cohen, M. (2025). Century-scale storage. *Harvard Law School Library Innovation Lab*. <https://lil.law.harvard.edu/century-scale-storage/>
- Risam, R. (2019). *New digital worlds: Postcolonial digital humanities in theory, praxis, and pedagogy*. Northwestern University Press. <https://doi.org/10.2307/j.ctv7tq4hg>
- Spetalnick, M., & Parraga, M. (2025, January 10). US and allies slap sanctions on Venezuela officials as Maduro inaugurated. *Reuters*. Accessed on January 17, 2025, from



<https://www.reuters.com/world/americas/us-slaps-new-sanctions-venezuela-officials-maduro-inaugurated-2025-01-10/>

UNESCO. (2015, November 17). *Recommendation concerning the preservation of, and access to, documentary heritage including in digital form*. Accessed on January 20, 2025, from <https://www.unesco.org/en/legal-affairs/recommendation-concerning-preservation-and-access-documentary-heritage-including-digital-form>

UNESCO. (2021a). *Documentary heritage at risk: Policy gaps in digital preservation*. Accessed on January 20, 2025, from [https://www.unesco.org/sites/default/files/documentary\\_heritage\\_at\\_risk\\_policy\\_gaps\\_in\\_digital\\_preservation\\_en.pdf](https://www.unesco.org/sites/default/files/documentary_heritage_at_risk_policy_gaps_in_digital_preservation_en.pdf)

UNESCO. (2021b). *Managing low-cost digitization projects in least developed countries and small island developing states: A manual*. Accessed on January 20, 2025, from <https://unesdoc.unesco.org/ark:/48223/pf0000380165>

UNESCO. (2023). *Information for all programme (IFAP) strategic plan 2023-2029*. Accessed on January 20, 2025, from <https://unesdoc.unesco.org/ark:/48223/pf0000386173/PDF/386173eng.pdf.multi>

UNESCO. (2024). *Advancing access to information and digital preservation: IFAP's impact on digitizing documentary heritage*. Accessed on January 20, 2025, from <https://www.unesco.org/en/articles/advancing-access-information-and-digital-preservation-ifaps-impact-digitizing-documentary-heritage>

United Nations High Commissioner for Refugees (2025). Venezuela situation. *Global Focus: UNHCR Operations Worldwide*. Accessed on January 20, 2025, from <https://reporting.unhcr.org/operational/situations/venezuela-situation>

Wikidata. (2024, December 26). *Wikidata: Introduction*. Accessed on January 20, 2025, from <https://www.wikidata.org/wiki/Wikidata:Introduction>